RAW SEQUENCE LISTING
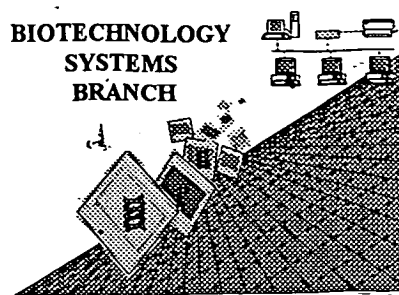# ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 09/831,621

Source: PCT/09

Date Processed by STIC: 10/11/2001

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:
1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**
FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO **REDUCE** ERRORED SEQUENCE LISTINGS, PLEASE USE THE <u>CHECKER VERSION 3.0 PROGRAM</u>, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW:

Checker Version 3.0

The Checker Version 3.0 application is a state-of the-art Windows based software program employing a logical and intuitive user-interface to check whether a sequence listing is in compliance with format and content rules. Checker Version 3.0 works for sequence listings generated for the original version of 37 CFR §§1.821 – 1.825 effective October 1, 1990 (old rules) and the revised version (new rules) effective July 1, 1998 as well as World Intellectual Property Organization (WIPO) Standard ST.25.
Checker Version 3.0 replaces the previous DOS-based version of Checker, and is Y2K-compliant. Checker allows public users to check sequence listings in Computer Readable form (CRF) before submitting them to the United States Patent and Trademark Office (USPTO).
Use of Checker prior to filing the sequence listing is expected to result in fewer errored sequence listings, thus saving time and money.

<u>Checker Version 3.0 can be down loaded from the USPTO website at the following address:</u>
http://www.uspto.gov/web/offices/pac/checker

RAW SEQUENCE LISTING   DATE: 10/11/2001
PATENT APPLICATION: US/09/831,621   TIME: 10:39:22

Input Set : A:\98,664-A.seq list.txt
Output Set: N:\CRF3\10112001\I831621.raw

```
    3 <110> APPLICANT: Baum, Peter
    4         DuBose, Robert
    5         Sims, John E
    6         Youakim, Adel
    7         Hasel, Karl W
    8         Hilbush, Brian S
   10 <120> TITLE OF INVENTION: Novel DNAs and Polypeptides
   12 <130> FILE REFERENCE: 98,664-A
   14 <140> CURRENT APPLICATION NUMBER: US/09/831,621
   15 <141> CURRENT FILING DATE: 1999-11-10
   17 <150> PRIOR APPLICATION NUMBER: 60/107821
   18 <151> PRIOR FILING DATE: 1998-11-10
   20 <160> NUMBER OF SEQ ID NOS: 33
   22 <170> SOFTWARE: PatentIn Ver. 2.1
```

ERRORED SEQUENCES

```
  125 <210> SEQ ID NO: 9
  126 <211> LENGTH: 180    170 shown
  127 <212> TYPE: DNA
  128 <213> ORGANISM: Homo sapiens
  130 <400> SEQUENCE: 9
  131 cggctgcctg cctttttttc tgatccagac cctcggcacc tgctacttac caactggaaa    60
  132 attttacgca tcccatgaag cccagataca caaaattcca ccccatgatc aagaatcctg   120
  133 ctccactaag aatggtgcta aagtaaaact agtttaataa gccctaaaaa    180   170
  145 <210> SEQ ID NO: 11
  146 <211> LENGTH: 569
  147 <212> TYPE: DNA
  148 <213> ORGANISM: Homo sapiens
  150 <400> SEQUENCE: 11
  151 tcaatcctgg gcggcgacaa gacagctcta gagatctgag cttcctccca atgctaaact    60
  152 gctttcatgc taattttctg actgtttact taccgggtaa gagcgatggg actgttttca   120
  153 ttggttggtt ctcacatact ctctgggaag tttgggttct cagggacacc tgctcctcag   180
  154 ctggggacca tggccatggc ccaccacctg cccttcagtg ttcaagcagg ggacatgcac   240
  155 cctttagtaa cctggagggg acccatcaca tgacaaccac cccaacgacc atcatcagga   300
  156 agccgctgcc tgactgagat atgcccccag gaggacaagg gagagtggat gctggaaaga   360
  157 cagggcaggg gaccatcacc agggaaagac ttcattcttc ggaggacatt gaacctgggg   420
  158 ctgggtctgt agtggagccg ctgtttcttc tcctgtatcc aactgttcta actcttgggc   480
  159 tttctccatt ttcagctctt tcttttcctg gccttctcat tgctgggtcc ttcaagcctc   540
  160 cgctctatgc ttccgncaat atattcttt                                      569
  182 <210> SEQ ID NO: 13
  183 <211> LENGTH: 618
  184 <212> TYPE: DNA
  185 <213> ORGANISM: Homo sapiens
  187 <400> SEQUENCE: 13
  188 cacgagctgt ctggttatta tacagacgca taactggagg tgggatccac acagctcaga    60
```

```
      189 acagctggat cttgctcagt ctctgccagg ggaagattcc ttggaggagg ccctgcagcg   120
      190 acatggaggg agctgctttg ctgagagtct ctgtcctctg catctggatg agtgcacttt   180
      191 tccttggtgt gggagtgagg gcagaggaag ctggagcgag ggtgcaacaa aacgttccaa   240
      192 gtgggacaga tactggagat cctcaaagta agcccctcgg tgactgggct gctggcacca   300
      193 tggacccaga gagcagtatc tttattgagg atgccattaa gtatttcaag gaaaaagtga   360
      194 gcacacagaa tctgctactc ctgctgactg ataatgaggc ctggaacgga ttcgtggctg   420
      195 ctgctgaact gcccaggaat gaggcagatg agctccgtaa agctctggac aaccttgcaa   480
      196 gacaaatgat catgaaagac aaaaactggc acgataaagg ccagcagtac agaaactggt   540
      197 ttctgaaaag agtttcctcg ggtgaaaaag taagcttgag gataacataa gaaagcttcc   600
E--> 198 gtgcccttgc aaatgggg                                                 618
```

item 9

Input Set : **A:\98,664-A.seq list.txt**
Output Set: **N:\CRF3\10112001\I831621.raw**

L:14 M:270 C: Current Application Number differs, Replaced Application Number
L:133 M:254 E: No. of Bases conflict, LENGTH:Input:180 Counted:170 SEQ:9
L:133 M:252 E: No. of Seq. differs, <211>LENGTH:Input:180 Found:170 SEQ:9
L:159 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:11
M:340 Repeated in SeqNo=11
L:198 M:340 E: (46) "n" or "Xaa" used: Feature required, for SEQ ID#:13

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 09/831, 621

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWAR

1 _____ Wrapped Nucleics
Wrapped Aminos

The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."

2 _____ Invalid Line Length

The rules require that a line not exceed 72 characters in length. This includes white spaces.

3 _____ Misaligned Amino
Numbering

The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.

4 _____ Non-ASCII

The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.

5 _____ Variable Length

Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.

6 _____ PatentIn 2.0
"bug"

A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.

7 _____ Skipped Sequences
(OLD RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i)   SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.

8 _____ Skipped Sequences
(NEW RULES)

Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000

9 _____ Use of n's or Xaa's
(NEW RULES)

Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.

10 _____ Invalid <213>
Response

Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence

11 _____ Use of <220>

Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32)   (Sec. 1.823 of Sequence Rules)

12 _____ PatentIn 2.0
"bug"

Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.

13 _____ Misuse of n

n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

AMC/MH – Biotechnology Systems Branch – 08/21/2001